

## Figure S1: Mis-alignment in SSAHA2 causes a false prediction in *TET2*.

IGV (Integrative Genomic Viewer) software visualizes alignment of next generation sequencing reads to the reference sequence.<sup>15</sup> In IGV, each sequence read is represented as a grey rectangle and the reference sequence is represented at the bottom. If there is base in a read that is different from the reference sequence, it is indicated with the corresponding letter. This figure shows the IGV output when analyzing the same set of reads with the BWA-SW (top) and with the SSAHA2 (bottom) algorithms for sequence alignment. Looking at the alignments generated by these two algorithms revealed that SSAHA2 alignment was incorrectly positioning 3 reads (as indicated by colored read names on the plot), causing a false variant call on chr4:106384366 and resulting false prediction of a non-synonymous coding mutation in the *TET2* gene.

